Net-Centric Information Management

Dr. Scott Renner
The MITRE Corporation
sar@mitre.org

Abstract

Information sharing is a key tenet of network-centric warfare (NCW). Information sharing succeeds when the right information is provided to the right people at the right time and place so that they can make the right decisions. This will not occur without an information management policy and process that is fitted to the needs of NCW – one that is flexible, seamless, and complete. In this paper we describe the essential architecture of a net-centric information management process, one that is based on the information and data management strategy of the US Air Force.

1. Introduction

NCW is about deriving combat power from distributed interacting entities with significantly improved access to information [1]. This improvement is derived in part from better communication networks, things built with cables, radio links, and TCP/IP, things that deliver data bits from one networking participant to another. A more important factor in this improvement is the participants' ability to find the data they need and to understand that data when they receive it. In the terms of the DoD Net-Centric Data Strategy [2], data must be visible, accessible, understandable, trusted, interoperable, and made available in response to user needs.

These things will not occur without an information management policy and process that is fitted to the needs of NCW, which imposes additional demands on the activities of information management. In this paper we describe the essential architecture of a net-centric information management process from the viewpoint of governance: what activities must be performed, and the roles/responsibilities of the people and organizations that perform them.

2. Information Management

Everyone agrees that information is essential to network-centric warfare. Not everyone agrees on the meaning of the term.

Within the DoD, the term "information" has two approved definitions: it is either data, in any medium or form, or it is the meaning a person assigns to that data [3]. In the first definition, "information" and "data" are synonyms; the term includes data of any kind, any sort of symbol or analog quantity that can have meaning. For example, the contents of a relational database, text file, spreadsheet, audio recording, image, and video segment are all different kinds of data and information. Paper documents are also both data and information. In the second definition, the term denotes a person's understanding of data

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUN 2005		**00-00-2005 to 00-00-2005**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Net-Centric Information Management	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Mitre Corporation,202 Burlington Road,Bedford,MA,01730**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**46**	

in the context of some decision. That is, information is data that makes a difference to a decision, data which "informs" the decider. We follow the first definition to describe the scope of information management; i.e., what it *is*. However, we use the second definition to describe its purpose; i.e., what it is *for*.

Information management (IM) is typically defined as "the planning, budgeting, manipulating, and controlling of information throughout its lifecycle" [4,5]. It may be understood as "a set of intentional activities which maximize the value of information in support of the objectives of the enterprise" [6]. These activities control information within the enterprise, from creation, through dissemination and use, to final disposition. Many of these activities are known and governed under a different name: data management, records management, content management, knowledge management, etc. All of these "managements" may properly be considered aspects of IM.

The value of information to the enterprise may have both external and internal origins. External value arises when information is produced for delivery to the outside world. Satellite images sold by commercial vendors and information maintained by government agencies solely to satisfy freedom-of-information requests and legislative inquiries fall into this category.

In this paper, we will focus on the internal value of information, which derives from its use in the decisions made within the enterprise. People make better decisions when they have the right information – data which they understand and which is relevant to the decision at hand. The internal purpose of information management is to supply the right information to the right people at the right time and place so that they can make the right decisions – the "five rights" of IM.

These "five rights" describe the purpose of the Air Force Information and Data Management Strategy [7]. That strategy sets three goals for Air Force information management activities: to ensure that the right data exists, is accessible, and is understood and discoverable. The first and most important step is to ensure that data is *accessible;* that is, made available by those who have it and deliverable to those who need it. The next step is to make the right data *discoverable and understandable.* Individuals and organizations must be able to obtain all the data they need, but to avoid the problem of data overload, it must be possible for them to receive only the data they need. Finally, the enterprise must take steps to ensure that the right data will *exist.* The enterprise must develop an understanding of current and anticipated information needs to drive the development and operation of its data resources, so that the data needed by a decider will be collected and made available somewhere in the enterprise.

Understanding these current and anticipated information needs is also an important part of *enterprise architecture.* Enterprise architectures exist to inform, guide, and constrain the decisions for the enterprise, especially those related to IT investment. Many of those decisions are related to IM activities; for example, planning and budgeting. We will observe more connections between architecture and IM in the discussion of common vocabularies (section 4.3).

3. Demands of Net-Centric Information Management

NCW places additional requirements on the activities of information management, in addition to the "five rights" that are the ordinary internal purpose of IM. These may be deduced from four predictions about the future NCW environment [8]:

- The network will include a very large number of participants. All will need to exchange information with some other participants.
- The friendly forces participating in the network will often be drawn from a coalition of sovereign organizations. These must be considered to be separate enterprises for many information management activities.
- Information technology (and the people who understand it) will become much less expensive and therefore widely available to adversaries. There will be little competitive advantage in IT *per se*. Advantage will come from knowing how to best employ the technology that will be available to everyone.
- Working out the best ways to employ IT will be an iterative process; a co-evolution of technology, doctrine, and organization. Organizations which make that iterative process go quickly will maximize their advantage.

We derive the following requirements from these predictions:

- *Flexibility and agility.* If information management activities are slow and rigid, then the enterprise coevolution process will be slowed, and opportunities for competitive advantage lost.
- *An end to pairwise arrangements.* At present, several aspects of information sharing are often arranged between individual producers and consumers, one pair at a time. These arrangements include semantics (does the consumer understand the producer's data?), data sharing implementations (exactly what data does the consumer obtain from the producer's system, and how?), and access control (is the consumer entitled to obtain the producer's data?). With many participants, there will be far too many pairs to arrange.
- *Success without recourse to a single central authority.* Coalition partners have no single governing authority during much of the information lifecycle. Even with a single authority, beyond a certain point of scale an enterprise becomes too large and diverse for centralized, top-down control. For both reasons, net-centric IM activities must make use of negotiation, influence, and competition, in addition to orders, authority, and top-down direction.

4. A Net-Centric Information Management Architecture

In this section we describe the key elements of net-centric IM and the relations between them. Our viewpoint will emphasize the organization and governance of information management activities. From this viewpoint we will summarize *what* activities must be

performed (omitting details about *how, where,* or *when* they will be done), and concentrate on the roles and responsibilities of *who* does them.

The four key elements of a net-centric IM architecture are: information owners, shared information spaces and their controlling authorities, common vocabularies within semantic communities, and the enterprise infrastructure for implementation. These are briefly defined below:

- *Information owners* are organizations that exercise authority over data. They are responsible for the production of data in the enterprise.
- *Shared information spaces* (or *infospaces*) are collections of data intended to suit the needs of different groups of consumers of data. Each has a controlling authority, who represents and often has authority over the consumers.
- *Common vocabularies* represent the shared understanding of terms held by the people in a *semantic community*. The detail in these vocabularies may range from simple dictionaries used in architecture descriptions, to the detailed data models and elements required for machine-to-machine data exchanges.
- The *enterprise implementation infrastructure* supports the information systems operated by data producers and consumers. It includes some core enterprise services required at runtime, and others needed during development.

In the rest of this section we describe these elements in more detail, and explain why each is necessary. We will show that combining any two results in failure to meet the special demands of net-centric information management.

4.1 Information owners and data producers

Information owners are organizations that control decisions about data: what data must be collected, how it will be represented and stored, how it will be validated, the required degree of accuracy, precision, and other quality factors, when it will be released, who is allowed to access and update it, how long it will be maintained, etc. They often delegate this responsibility to subordinate organizations. Data producers are those information owners at the bottom of the delegation chain, those finally accountable for the data. Their accountability includes certain responsibilities sometimes assigned by public law and regulations; for example, requiring its protection against unauthorized access. Figure 1 shows a notional arrangement of information owners into a tree, with data producers at the leaf nodes.

Information owners acquire and operate information systems to carry out their ownership responsibilities. An information system is not a data producer; it is built on behalf of a data producer, and operated by a data producer. (However, it is often useful to talk about the "producing system".) Individual data entry operators are not data producers; they are people who work for a data producer.

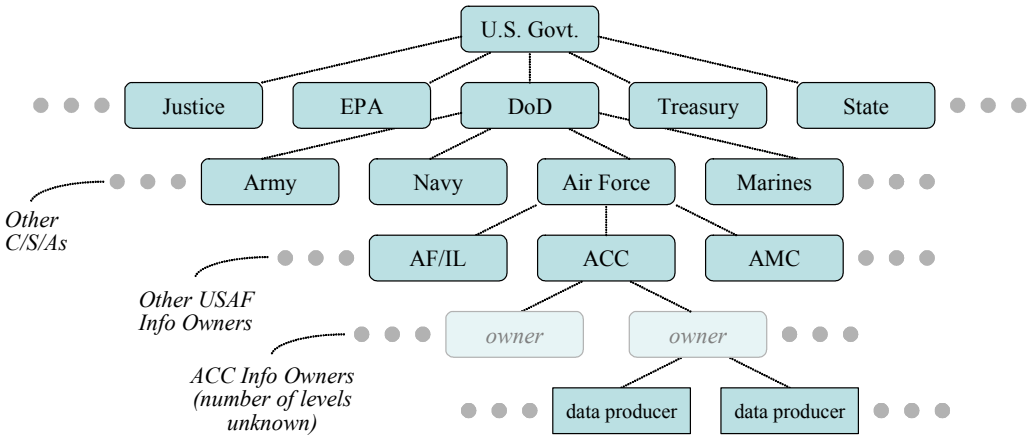


Figure 1: A Notional Information-Owner Tree

Figure 1 illustrates why an enterprise can have data ownership problems even though "all data belongs to the enterprise". As a practical necessity, ownership authority and responsibility must be delegated to low levels in the tree. Information sharing arrangements frequently cut across the branches of the tree. However, data consumers usually have little influence on information owners in a completely different subtree. This tends to make cross-organizational data sharing more difficult.

Every enterprise already has people who are making the decisions and doing the work that information owners make and do. They are already accountable within the organizational hierarchy. Governance of information owners should follow the existing structure. What is required is a process that will make information ownership decisions visible to any interested party, while efficiently including and resolving the needs of all stakeholders, so that these ownership decisions are made for the benefit of the entire enterprise.

4.2 Shared information spaces

A shared information space is a collection of data intended to suit the needs of a group of consumers. Data producers post data to one or more infospaces; data consumers pull the data they need from one or more infospaces. The defining aspects of an infospace are the data content, the governance process, the infospace controller, and the consumers subject to that authority.

Infospace governance concerns decisions about the infospace contents and the consumers who access that content. The governing authority controlling these decisions may vary. Some infospaces have a single governing authority, typically the commander for whom all the consumers are working. In others, several relatively autonomous organizations establish through consensus the authority for their shared information space.

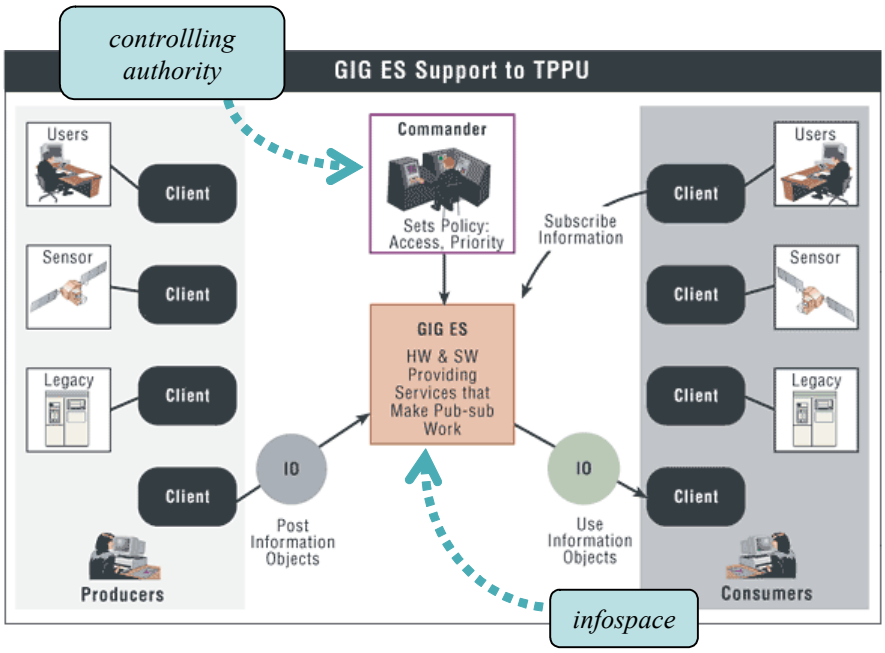


Figure 2: An Infospace and Controlling Authority in the TPPU Paradigm

---

Each infospace has an executive agent, who typically exercises some form of control over some of the IT resources that implement the infospace (systems, networks, etc.). This infospace controller executes the decisions of the infospace authority: control over which producers are allowed to post to the infospace, what kinds of data they may post, the frequency at which they may post data, which of these sources will be authoritative for the infospace consumers, and how the infospace data should be organized for navigation and discovery. The infospace controller may track and validate the information needs of the infospace consumers, and may search out data producers to satisfy these needs, possibly negotiating with producers to obtain new kinds of data not yet collected or maintained.

The infospace controller must also enforce the access control policies set by the infospace authority. This includes establishing the roles for role-based access, their privileges, and the assignment of roles to individuals. Finally, the infospace controller establishes priorities for consumers and arbitrates their conflicting quality-of-service demands. Figure 2 contains the GIG Enterprise Services illustration of the new *Task Post Process Use (TPPU)* paradigm [9], which also serves as a good illustration of a shared information space.

The infospace controller is essential to the ending of pairwise access control arrangements in two ways. First, the data producers no longer need to determine the access privileges of each individual consumer. Instead, infospace controllers assign roles to consumers and assign privileges to roles. Data producers decide whether they will post their data to the infospace, given the controller's declared access policy. If there are N producers and M consumers, the decision workload is thereby reduced from N x M

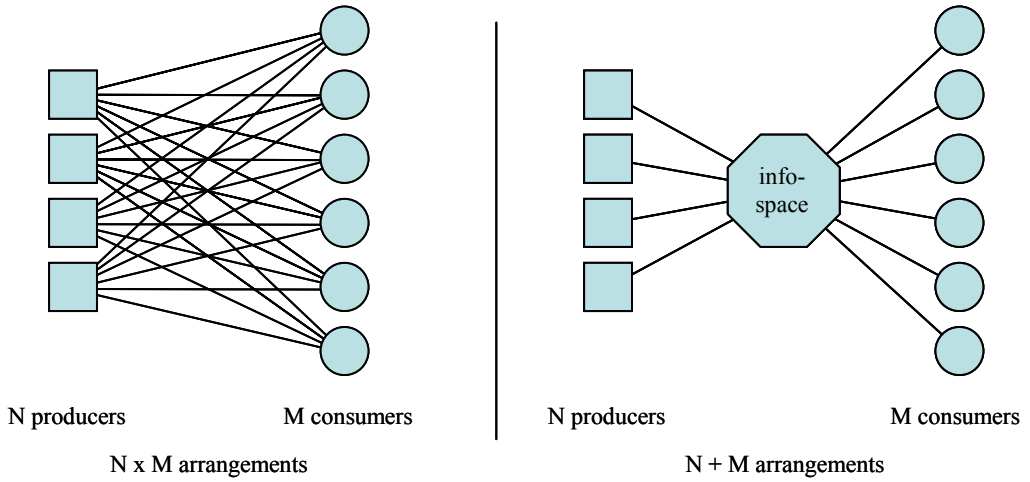


Figure 3: Access Control With and Without Infospace Controller

producer-consumer pairs to something much smaller: N producers deciding to post, plus M consumer privilege decisions made by the infospace controller. This is illustrated in Figure 3.

Second, the infospace controller is an entity able to accept a transfer of accountability for the posted data. Data producers are often legally accountable for controlling access to their data. Appropriate policy and regulations can establish that if a data producer posts to an infospace, and the controller fails to enforce his declared access control policy, then the controller is held responsible, not the producer. Such an arrangement is the only alternative to the pairwise approval of each consumer by each data producer.

The governance question for shared information spaces is: How is the controlling authority established, and what is its relation to the infospace consumers? In some cases the enterprise will follow the existing lines of combatant command authority or its equivalent in the business mission area, creating infospaces to be governed by the NCA, the CoCOMs, or perhaps lower in the command structure. In other cases, the enterprise will create more collaborative infospaces established by consensus among autonomous organizations, and governed through a process satisfactory to all. When creating infospaces, the following principle will hold true: distinct groups of consumers for whom no single infospace controller can be established (by command or consensus) require separate infospaces. Over time, examples of successful infospace governance will turn into "templates" suitable for reuse.

Some arrangements will be made by negotiation between producers and infospace controllers. They will need to form (and record) at least the following two kinds of agreements:

- Availability agreements between a producer and an infospace controller. These describe the conditions under which the infospace controller may rely on the

availability of information from the producer. The terms may include effective dates (beginning and ending), measures of data quality, measures of performance, etc.

- Access control agreements between a producer and an infospace controller. These describe access control guarantees made by the infospace controller to the data producer, and/or the terms under which accountability for access control enforcement is transferred from producer to infospace controller.

We have said nothing about the implementation of shared information spaces, because implementation has little or no impact on governance. Implementation is not a defining aspect – there are several possible "styles" of infospace implementation. These are discussed in section 4.4.

4.3 Semantic communities and common vocabularies

Producers and consumers of data (and the people who build their information systems) must have a shared understanding of what the data means. If their understandings are incompatible, the result will be mistakes, failures of interoperability, and suboptimal decisions, caused when one person (or automated system) misinterprets the data provided by another. Our goal is to get the right information to the right person at the right time, and we cannot succeed unless all of the people involved have a compatible understanding of what the right data *is*, and what it *means*. These people comprise a semantic community. Their shared understanding is represented in the community's common vocabulary.

Semantic communities are much the same as the *communities of interest (COIs)* described in the DoD Net-Centric Data Strategy. There, COIs are defined as "collaborative groups of users who must exchange information in pursuit of their shared goals, interests, missions, or business processes and who therefore must have shared vocabulary for the information they exchange." The COI term has since acquired a number of additional meanings, which is why we do not use it in this paper.

The people in a semantic community often have different uses for their common vocabulary. For example, some may be concerned with architectural descriptions of information exchange, others with discovery metadata "tags" and data asset catalogs, still others with implementing machine-level data sharing. As a consequence, community vocabularies may not be alike in their level of detail. However, the fundamental purpose is still always to establish a compatible understanding of terms.

It is almost always necessary to record a community vocabulary in some tangible format. Communities need this documentation to help teach new members what they need to know about the common vocabulary. They need it to remind current members of what they need to know. They need it to support tools that help the members do their jobs – which could be describing desired information flows, or discovering new information, or implementing application-level exchanges, or understanding the information they receive each day. The value of the recorded vocabulary lies in these activities – vocabulary is useless documentation unless it is comprehended by the community. For this reason, we

argue that the creation and use of a common vocabulary is a knowledge management problem [10].

We would like to have a single formalism for recording common vocabularies, but at present there is none suitable for all purposes. Instead, there are several alternatives, each aimed at a slightly different purpose; these include: ISO 11179 data element definitions, IDEF1X data models, UML, XML Schema, and semantic web languages (e.g. RDF, OWL).

We would very much like to have a single common vocabulary for the whole enterprise, but this is not possible beyond a certain point of size and complexity. The other extreme, with an enormous number of vocabularies negotiated between every pair of information sharers, is also unworkable. Semantic communities are the means of searching for the optimum tradeoff. We want to establish communities in a way that maximizes the value of exchangeable information while minimizing the cost of developing their common vocabularies.

At present, choosing the proper scope of a particular semantic community and its vocabulary is more of an art than a science. Some suggestions are offered in [10]. However, the overall pattern may be predicted: We will see a small number of broad and shallow vocabularies, understood by many people, while containing only a few definitions. We will also see a larger number of narrow and deep vocabularies, understood by a few people, while containing many more definitions. These will be arranged in a hierarchy, in which the lower vocabularies extend and specialize the higher. This pattern is illustrated in Figure 4. Broad-and-shallow vocabularies are sometimes known as "loose connectors", because they can facilitate a useful degree of information over the *intersection* of information needs, without requiring the work of establishing a single vocabulary for the *union* of those needs.

We observe that common vocabularies can be built by consensus. In the commercial world, this is the usual case; e.g. HL7 [11], RosettaNet [12], and DMTF [13]. It is

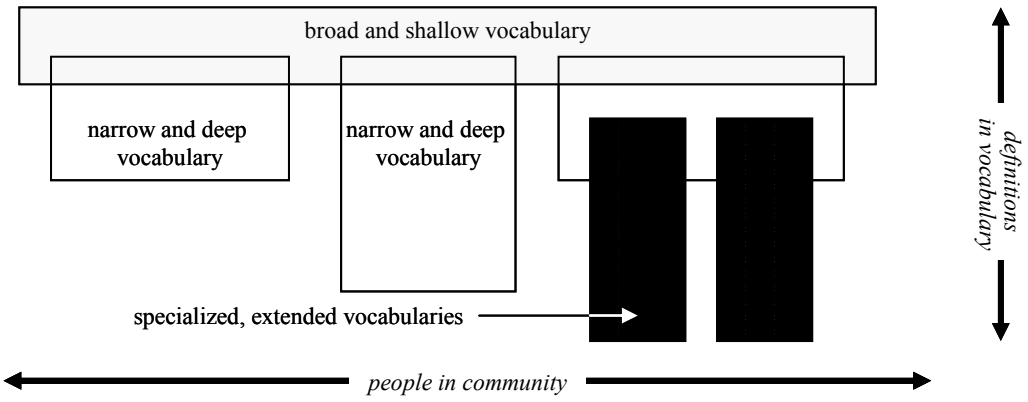


Figure 4: Pattern of Vocabulary Relationships

always possible for the community members to establish consensus definitions of the things they need to talk about. It is not necessary for the members to use those definitions internally; a data mediation service can resolve any differences in name, structure, or representation, once the people have agreed that the data meanings are compatible. If there is value in creating a common vocabulary, communities will often form bottom-up to do so.

In contrast, information owners and shared information spaces typically cannot be governed by consensus, because the choices involve real resources – does system A collect facts X, Y, and Z, or not? Can consumer B access these facts, or not? – the disputes can't be mediated away, and there may not be any choice that satisfies everyone.

The net-centric demand for flexibility, together with the present inability to determine optimum community boundaries in advance, means that the enterprise needs a governance process which allows some semantic communities to be created with top-down authority, and to form with bottom-up spontaneity. The rules should therefore allow any group of people to form a semantic community and develop their common vocabulary. However, the rules should control the *naming* of all semantic communities, perhaps reserving "blocks" of prominent names for the most important communities, those formed top-down, with a charter from an appropriate authority.[1] For example, there is probably no problem with a group of users and developers deciding to form an "air mission planning" community to define a common vocabulary for their systems. On the other hand, allowing a small and parochial community to name their work the "Joint Warfighting" common vocabulary is a bad idea that can only cause confusion.

Data producers and infospace controllers are not necessarily bound to follow the consensus definitions of any community. For one thing, there will often be multiple communities and multiple vocabularies capable of defining the information they produce or consume. Instead, data producers choose the vocabulary they will use internally, and the vocabularies they will support in external interfaces. Likewise, each infospace controller chooses the vocabularies that define the information his infospace contains. Vocabularies that add value will be chosen naturally. Also, data producers and infospace controllers will face pressure to adopt the vocabularies of the "official", chartered, important communities.

4.4 Enterprise implementation infrastructure

The primary reason to consider the implementation infrastructure as a separate element in the architecture is to show that the infospace controller is not solely responsible for the information sharing implementations. Without this architecture element, producers and consumers typically examine the architecture and imagine that information sharing will be delivered to them by an intermediate party with no effort on their part. In fact, both producing systems and consuming systems will typically bear some of the implementation burden.

---

[1] This might resemble the convention for creating USENET newsgroups. Anyone can create a newsgroup in the *alt.** hierarchy. There is an approval process for creating newsgroups in the main hierarchies.

The infospace controller is responsible for choosing the implementation style of the shared information space, and often for providing some of the implementation. Several styles are possible:

- a single physical database (or enterprise data warehouse)
- a distributed database
- a federation of semi-autonomous databases
- a peer-to-peer data sharing network
- a publish-and-subscribe message-passing network (with or without persistence of the shared data)

Information sharing implementation will depend on a small number of services that are available across the enterprise. The *DoD Metadata Registry* [14] is an example of a service needed during system development. Infrastructure services like the *GIG Core Enterprise Services* [15] will be required at runtime.

4.5 Necessity of the architecture elements

In this section we contend that the four IM architecture elements are necessary. First, we wish to deal with a seeming counterexample from recent history. The DoD data administration program conducted through the 1990s has been adjudged an overall failure [2], but did produce some local pockets of success. None of those success stories involved a careful separation of information owners, shared information spaces, and common vocabularies.

An examination of those success stories explains how such success is sometimes possible. In every case we find that the information owners and the data consumers were all subject to the same authority, in a cohesive enterprise with few external interfaces and of limited size and complexity. In such cases effective central governance is feasible, and separate governance of owners, infospaces, and vocabularies unnecessary. This is a good thing whenever it is possible. But it is not always possible – and net-centric IM must cope with all situations, not just those that are easy.

NCW will always involve situations where the information sharers are *not* all subject to the same effective central authority; this is in fact usually the case. Such situations are described in [16] and illustrated in Figure 5 on the following page. They include:

- Data producers that post to more than one shared information space
- Data producers that must understand more than one vocabulary
- Common vocabularies that are used in more than one shared information space
- Shared information spaces that include data defined in separate vocabularies.

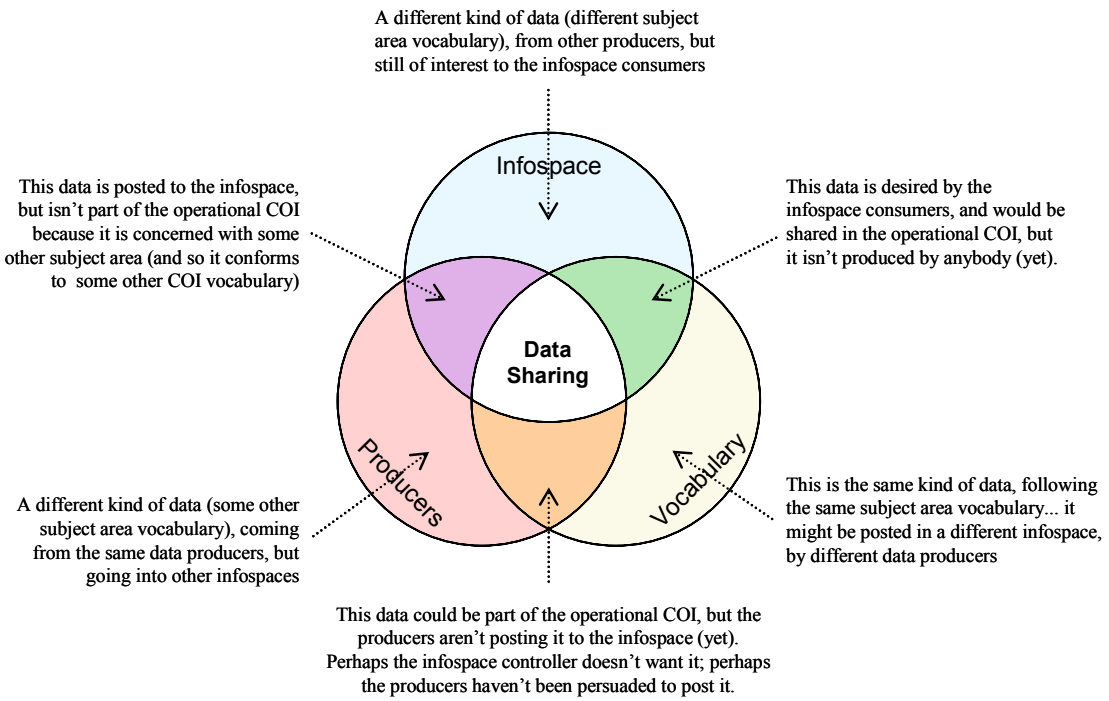


Figure 5: Possible Overlaps Between Architecture Elements

We will demonstrate the necessity of separate governance by considering and rejecting the three possible pairings below.

4.5.1 Shared vocabulary and information owners. These can be combined and governed together only when all of the people who need a common vocabulary are reporting to the same information owner. Usually they aren't. There are then three logical possibilities:

- One possibility is to ascend the information-owner hierarchy, until an owner that spans the semantic community is reached. Such owners are typically *very* high in the information-owner tree, and the corresponding community is always too large and diverse for a single vocabulary. For example, the answer might be to say "all data is owned by the DoD, and that community includes everyone who needs to share data ", but then the rest of the answer must be, "and there will be a single comprehensive vocabulary for the whole DoD". This has been attempted, without success.[2]
- It is also possible is to form a separate community for each owner. However, the separate communities will inevitably develop different vocabularies for the same kind of information. The result is the creation of information stovepipe walls along the boundaries of the information owners. Basically, this situation is the *status quo.* It is not the situation we desire.

[2] Even if it did succeed, what about the French Army? Or United Airlines? These will never be subordinate to the DoD, and yet there must be some shared vocabulary so that the DoD can share data with them.

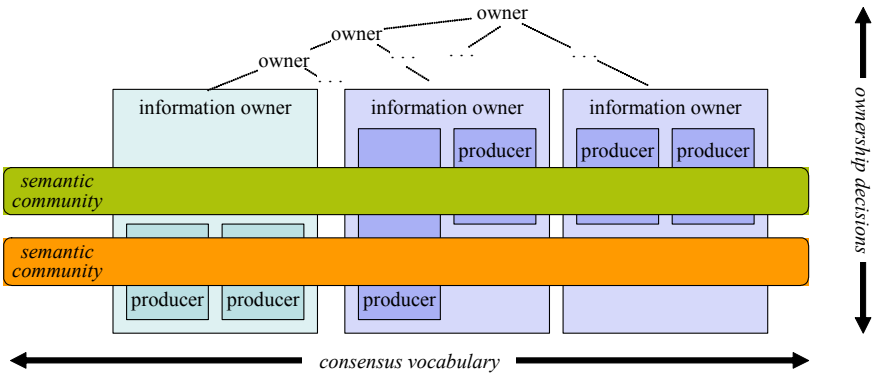


Figure 6: Semantic Communities and Information Owners

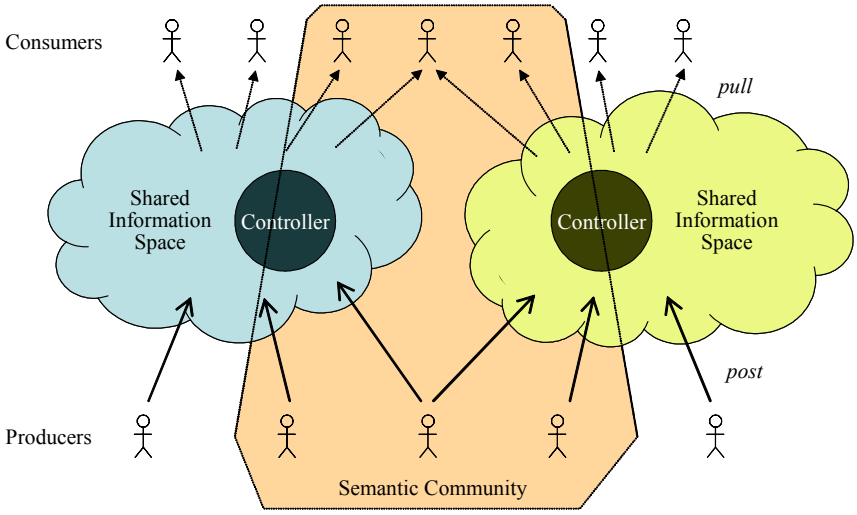


Figure 7: Semantic Communities Span Infospaces

- Finally, there can be communities that span information owners, and information owners that belong to multiple communities. None is a subset of another, so they must be separate. This situation is illustrated above in Figure 6.

4.5.2 Shared vocabulary and shared information spaces. These can be combined and governed together when the people who need a common vocabulary for sharing a particular kind of information are also consumers of a single infospace. This will not always be possible. It may be ruled out without having to know how many infospaces the DoD will need, or who will control them. For example, it is likely that the combatant commands will control separate infospaces. It is certain that those commands will be interested in some of the same kind of information; e.g. SIGINT. A separate community for each infospace is highly undesirable – they will develop different vocabularies, producers must then understand all of them, and consumers must learn a new vocabulary

when they are reassigned. Therefore, semantic communities and vocabularies must sometimes span multiple infospaces. This is depicted in Figure 7.

4.5.3 Information owners and shared information spaces. These can be combined and governed together when the infospace controller also controls all the data producers, or vice versa. This will not always be possible. Consumers often need data from producers in completely separate organizations. Data producers often post information of interest to consumers in distinct organizations.

5. Conclusion

Information management is a set of intentional activities which maximize the value of information in support of the objectives of the enterprise. NCW places additional demands on the ordinary goals of information management: a strong need for flexibility and agility, an end to pairwise arrangements, and the ability to function without a single central authority. These demands can be met by an information management architecture in which IM activities are performed and governed by distinct information owners, shared information space controllers, and semantic communities, all of which rely on the enterprise implementation infrastructure. This separation is necessary to accommodate the limited autonomy that is inevitable in the NCW enterprise, allowing for negotiation and competition in addition to top-down direction. Adopting information management policy and procedures which follow this architecture offers the best chance of satisfying the information sharing needs of NCW.

References

[1] D. Alberts, J. Gartstka, and F. Stein, *Network Centric Warfare: Developing and Leveraging Information Superiority*, 2nd ed., CCRP Publication Series, 1999. http://www.dodccrp.org/publications/pdf/Alberts_NCW.pdf

[2] DoD Chief Information Officer, *DoD Net-Centric Data Strategy,* March 2003. http://www.defenselink.mil/nii/org/cio/doc/Net-Centric-Data-Strategy-2003-05-092.pdf

[3] Joint Publication 1-02, *DoD Dictionary of Military and Associated Terms,* 2001. http://www.dtic.mil/doctrine/jel/new_pubs/jp1_02.pdf

[4] Office of Management and Budget, *Circular A-130: Management of Information Resources*, Nov. 2000. http://www.whitehouse.gov/omb/circulars/a130/a130trans4.pdf

[5] DoD Directive 8000.1, *Management of DoD Information Resources and Information Technology*, Feb. 2002. http://www.dtic.mil/whs/directives/corres/pdf/d80001wch1_022702/d80001p.pdf

[6] M. Linderman, J. Brichacek, S. Haines, D. Ouellet, B. Siegel, G. Chase, and J. O'May, "A Reference Model for Information Management to Support Coalition Information Sharing Needs", in *Proc. 10th Int. C2 Research and Technology Symposium,* McLean, VA, June 2005.

[7] U.S. Air Force, *Air Force Information and Data Management Strategy Policy,* Mar. 2004.

[8] S. Renner, "Building Information Systems For Network-Centric Warfare", in *Proc. 8th Int. C2 Research and Technology Symposium*, Washington, DC, June 2003.

[9] Global Information Grid (GIG) Enterprise Services (GES) Portal. http://ges.dod.mil/tppu.html

[10] S. Renner, "A Community of Interest Approach to Data Interoperability", in *Proc. Federal Database Colloquium*, San Diego, 2001.

[11] *Health Level Seven,* http://www.hl7.org.

[12] *RosettaNet,* http://www.rosettanet.org.

[13] *Distributed Management Task Force,* http://www.dmtf.org.

[14] DoD Metadata Registry and Clearinghouse. http://.metadata.dod.mil.

[15] DoD Chief Information Officer, *Global Information Grid Core Enterprise Services Strategy,* 2003. http://www.defenselink.mil/nii/org/cio/doc/GIG_ES_Core_Enterprise_Services_Strategy_V1-1a.pdf

[16] S. Renner, D. Hebert, S. Rainier, J.Wilson, "COI Handbook: Practical Guidance for Communities of Interest (COIs) Implementing the DoD Net-Centric Data Strategy", MITRE Technical Report, McLean VA, 2004.

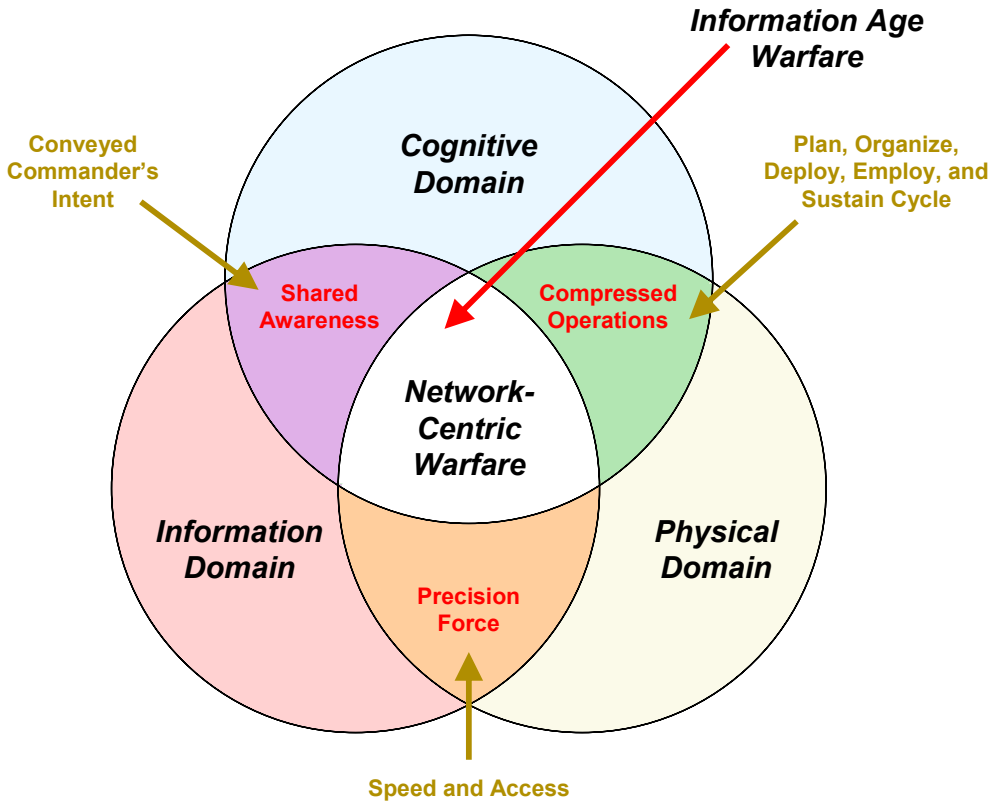
Net-Centric Information Management

Dr. Scott Renner
sar@mitre.org
13 June 2005

MITRE

Net-Centric Warfare

- Seamless interoperability
  - The network is only the beginning!
- Permits sharing of
  - Information
  - Situational awareness
  - Commander's intent
- Leading to
  - Speed of command
  - Self-synchronization
  - Enemy lock-out
- Producing increased combat power



NCW: Creating A Decisive Advantage
ASD/NII, Office of Force Transformation

Net-Centric Information Management

MITRE

The Fortune Teller's Predictions

❶ **Seamless network connectivity**

MITRE

The Fortune Teller's Predictions

1. Seamless network connectivity
2. **Very many network participants**

Expeditionary Sensor Grid
>10K distributed,
networked sensors

MITRE

The Fortune Teller's Predictions

1. Seamless network connectivity
2. Very many network participants
3. **Bandwidth limits at the sharp end**

MITRE

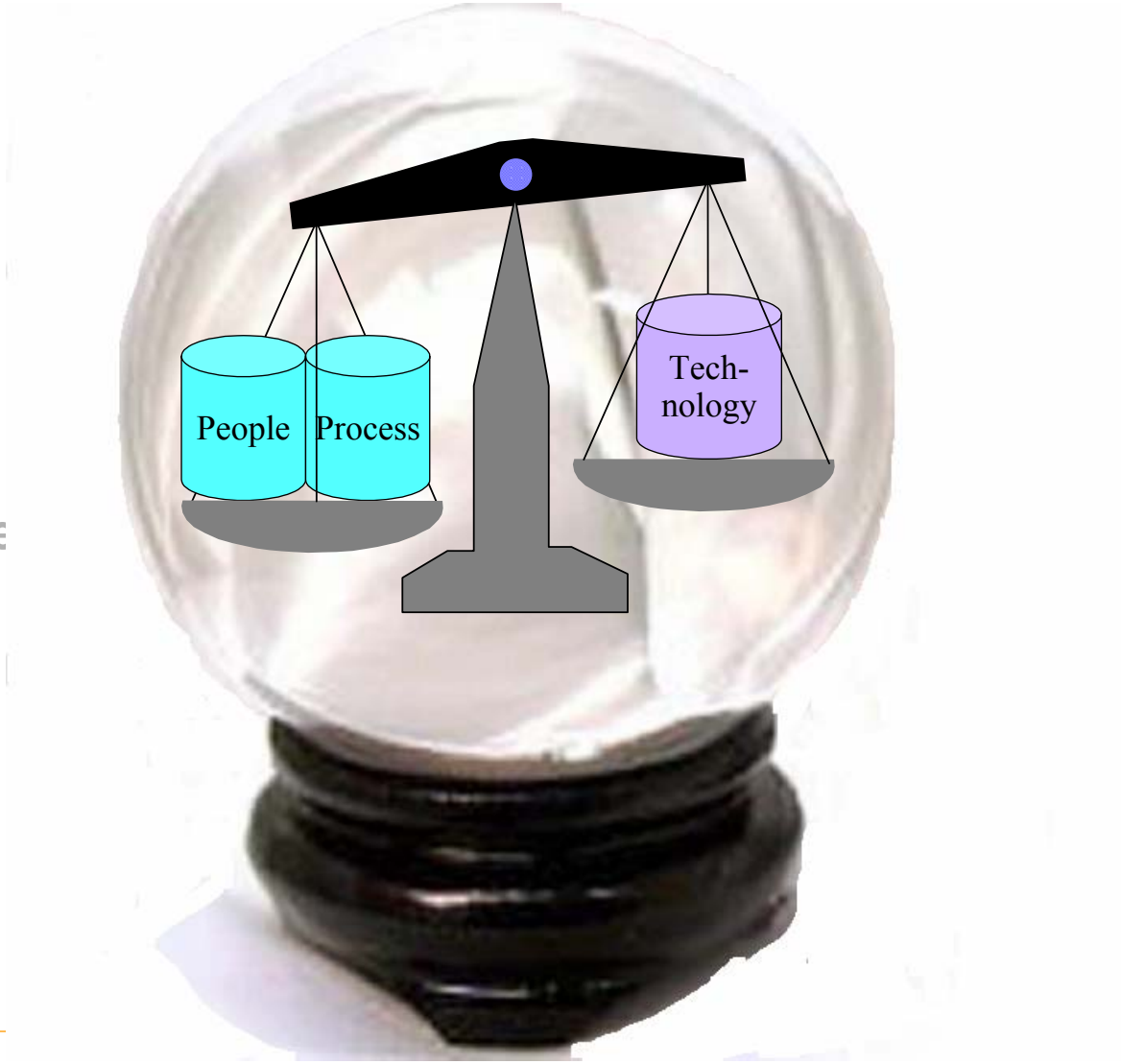
The Fortune Teller's Predictions

1. Seamless network connectivity
2. Very many network participants
3. Bandwidth limits at the sharp end
4. **Information assurance still crucial**

MITRE

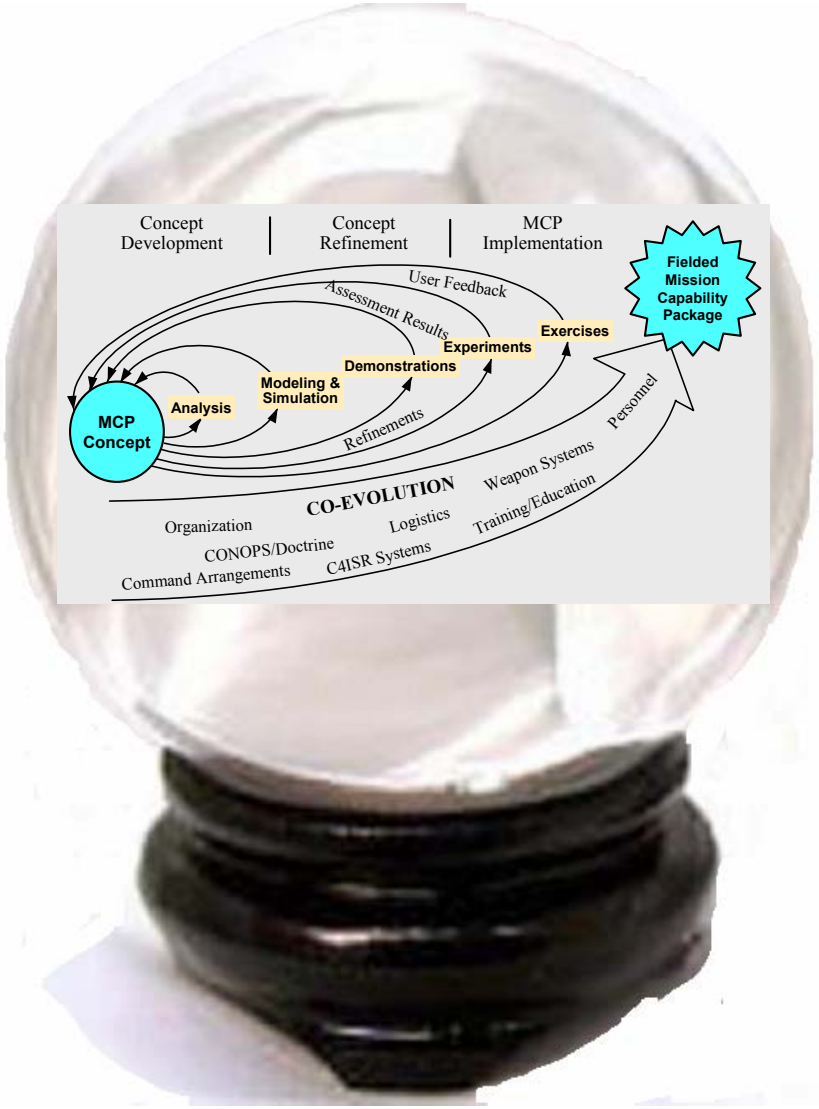
The Fortune Teller's Predictions

1. Seamless network connectivity
2. Very many network participants
3. Bandwidth limits at the sharp end
4. Information assurance still crucial
5. **Advantage comes from best use of IT**



The Fortune Teller's Predictions

1. Seamless network connectivity
2. Very many network participants
3. Bandwidth limits at the sharp end
4. Information assurance still crucial
5. Advantage comes from best use of IT
6. **Flexibility essential for quick coevolution**



Implications for Net-Centric IM

1. Seamless network connectivity
2. Very many network participants
3. Bandwidth limits at the sharp end
4. Information assurance still crucial
5. Advantage comes from best use of IT
6. **Flexibility essential for quick coevolution**

MITRE

Implications For Net-Centric IM

❶ Seamless network connectivity

❷ Very many network participants

❸ Bandwidth limits at the sharp end

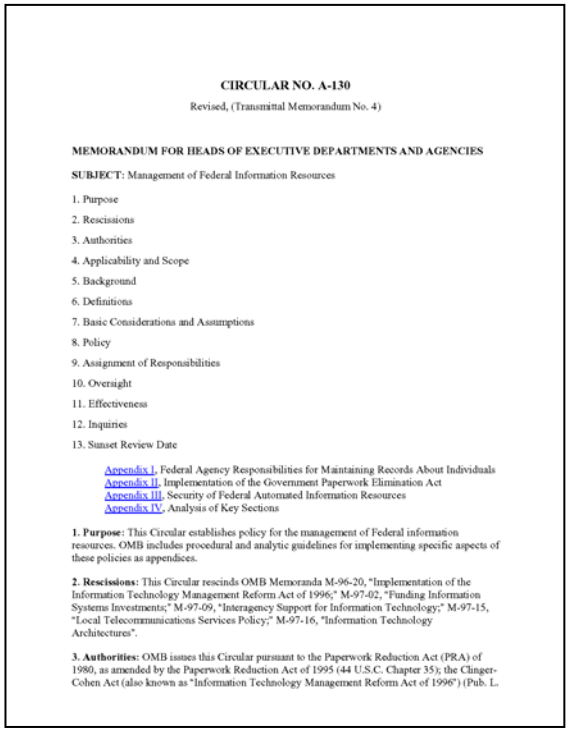
❹ Information assurance still crucial

❺ Advantage comes from best use of IT

❻ Flexibility essential for quick coevolution

- End of pairwise arrangements
  - For semantics
  - For information flows
  - For access control
- Shared semantics
  - Aligned to doctrine, training
  - Not one big data model
  - Not everyone does own thing
- Pedigree
  - Trust
  - Authoritative source
- Information preplanning
- Accountable data owners

Information Management, Officially Defined



CIRCULAR NO. A-130

Revised, (Transmittal Memorandum No. 4)

MEMORANDUM FOR HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

SUBJECT: Management of Federal Information Resources

1. Purpose
2. Rescissions
3. Authorities
4. Applicability and Scope
5. Background
6. Definitions
7. Basic Considerations and Assumptions
8. Policy
9. Assignment of Responsibilities
10. Oversight
11. Effectiveness
12. Inquiries
13. Sunset Review Date

Appendix I, Federal Agency Responsibilities for Maintaining Records About Individuals
Appendix II, Implementation of the Government Paperwork Elimination Act
Appendix III, Security of Federal Automated Information Resources
Appendix IV, Analysis of Key Sections

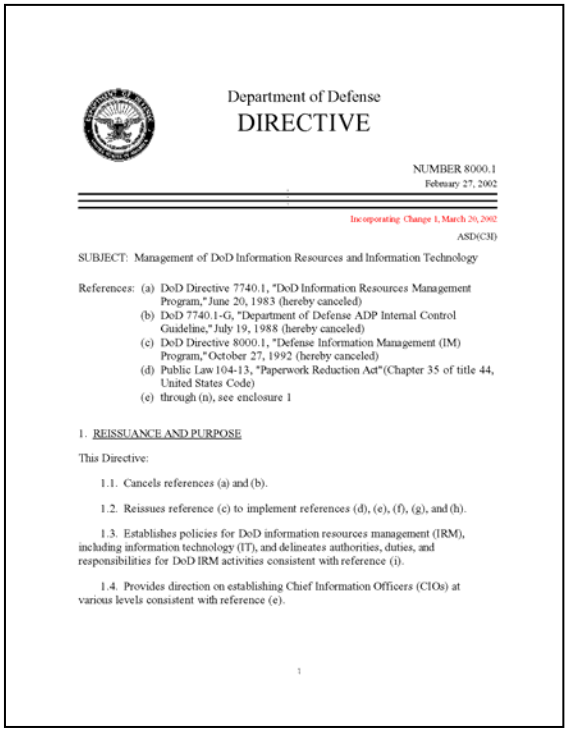
1. Purpose: This Circular establishes policy for the management of Federal information resources. OMB includes procedural and analytic guidelines for implementing specific aspects of these policies as appendices.

2. Rescissions: This Circular rescinds OMB Memoranda M-96-20, "Implementation of the Information Technology Management Reform Act of 1996;" M-97-02, "Funding Information Systems Investments;" M-97-09, "Interagency Support for Information Technology;" M-97-15, "Local Telecommunications Services Policy;" M-97-16, "Information Technology Architectures".

3. Authorities: OMB issues this Circular pursuant to the Paperwork Reduction Act (PRA) of 1980, as amended by the Paperwork Reduction Act of 1995 (44 U.S.C. Chapter 35); the Clinger-Cohen Act (also known as "Information Technology Management Reform Act of 1996") (Pub. L.

OMB Circular A-130
Management of Federal Information Resources

Requires Federal agencies to develop enterprise architectures



Department of Defense
DIRECTIVE

NUMBER 8000.1
February 27, 2002

Incorporating Change 1, March 20, 2002

ASD(C3I)

SUBJECT: Management of DoD Information Resources and Information Technology

References: (a) DoD Directive 7740.1, "DoD Information Resources Management Program," June 20, 1983 (hereby canceled)
(b) DoD 7740.1-G, "Department of Defense ADP Internal Control Guideline," July 19, 1988 (hereby canceled)
(c) DoD Directive 8000.1, "Defense Information Management (IM) Program," October 27, 1992 (hereby canceled)
(d) Public Law 104-13, "Paperwork Reduction Act" (Chapter 35 of title 44, United States Code)
(e) through (n), see enclosure 1

1. REISSUANCE AND PURPOSE

This Directive:

1.1. Cancels references (a) and (b).

1.2. Reissues reference (c) to implement references (d), (e), (f), (g), and (h).

1.3. Establishes policies for DoD information resources management (IRM), including information technology (IT), and delineates authorities, duties, and responsibilities for DoD IRM activities consistent with reference (i).

1.4. Provides direction on establishing Chief Information Officers (CIOs) at various levels consistent with reference (e).

DoD Directive 8000.1
Management of DoD Information Resources

Establishes the DoD and Component CIOs

Information Management, Officially Defined

OMB Circular A-130 and DoD Directive 8000.1 boil down to the same definitions

- **Information = data**
- **Information management :**

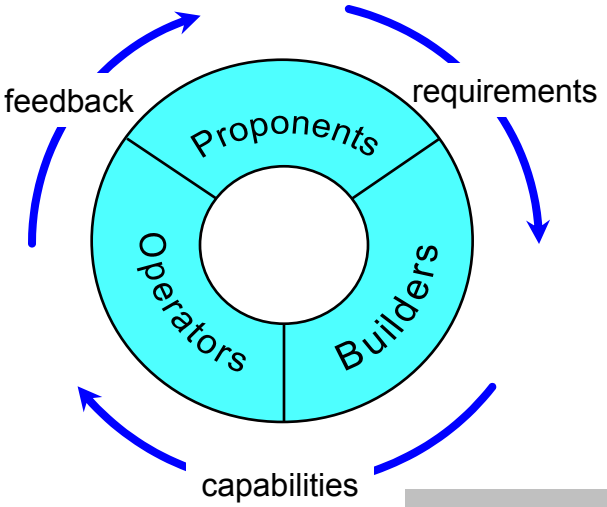
Planning, budgeting, manipulating, controlling of information throughout its lifecycle: creation / collection, processing, dissemination, use, storage, and disposition

- **These IM tasks apply to all information / data**
  - **Combat operations, combat support, and business data**
  - **Data for machine processing and human presentation**
  - **Not just records, documents, content**

Information Management, Officially Defined

Information Management:
Planning, budgeting, manipulating, controlling of information throughout its lifecycle: creation / collection, processing, dissemination, use, storage, and disposition

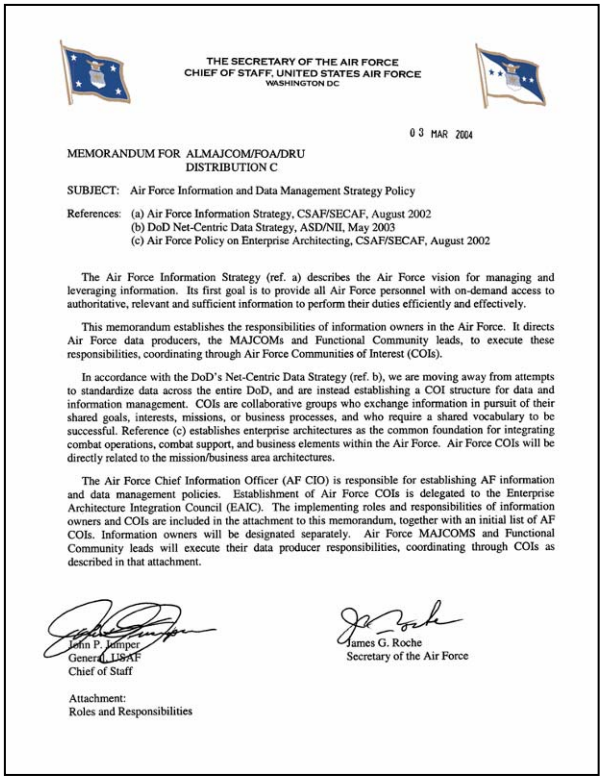
- **Note that proponents, builders, users all have a role in IM**
  - **Proponents: define the processes to be supported**
  - **Builders: construct systems to support process**
  - **Operators / users: employ system to conduct process**



Require, Build, Operate (RBO) Loop

AF Information and Data Management Strategy

- **Purpose of the strategy is to provide people the right data at the right time and place so that they can make the right decisions**
- **Goals: Make sure the right data**
  - **Exists**
  - **Can be discovered**
  - **Can be understood**
  - **Is accessible**
- **Policy memo defines roles and responsibilities**
  - **Information owners**
  - **Communities of Interest (COIs)**



THE SECRETARY OF THE AIR FORCE
CHIEF OF STAFF, UNITED STATES AIR FORCE
WASHINGTON DC

03 MAR 2004

MEMORANDUM FOR ALMAJCOM/FOA/DRU
DISTRIBUTION C

SUBJECT: Air Force Information and Data Management Strategy Policy

References: (a) Air Force Information Strategy, CSAF/SECAF, August 2002
(b) DoD Net-Centric Data Strategy, ASD/NII, May 2003
(c) Air Force Policy on Enterprise Architecting, CSAF/SECAF, August 2002

The Air Force Information Strategy (ref. a) describes the Air Force vision for managing and leveraging information. Its first goal is to provide all Air Force personnel with on-demand access to authoritative, relevant and sufficient information to perform their duties efficiently and effectively.

This memorandum establishes the responsibilities of information owners in the Air Force. It directs Air Force data producers, the MAJCOMs and Functional Community leads, to execute these responsibilities, coordinating through Air Force Communities of Interest (COIs).

In accordance with the DoD's Net-Centric Data Strategy (ref. b), we are moving away from attempts to standardize data across the entire DoD, and are instead establishing a COI structure for data and information management. COIs are collaborative groups who exchange information in pursuit of their shared goals, interests, missions, or business processes, and who require a shared vocabulary to be successful. Reference (c) establishes enterprise architectures as the common foundation for integrating combat operations, combat support, and business elements within the Air Force. Air Force COIs will be directly related to the mission/business area architectures.

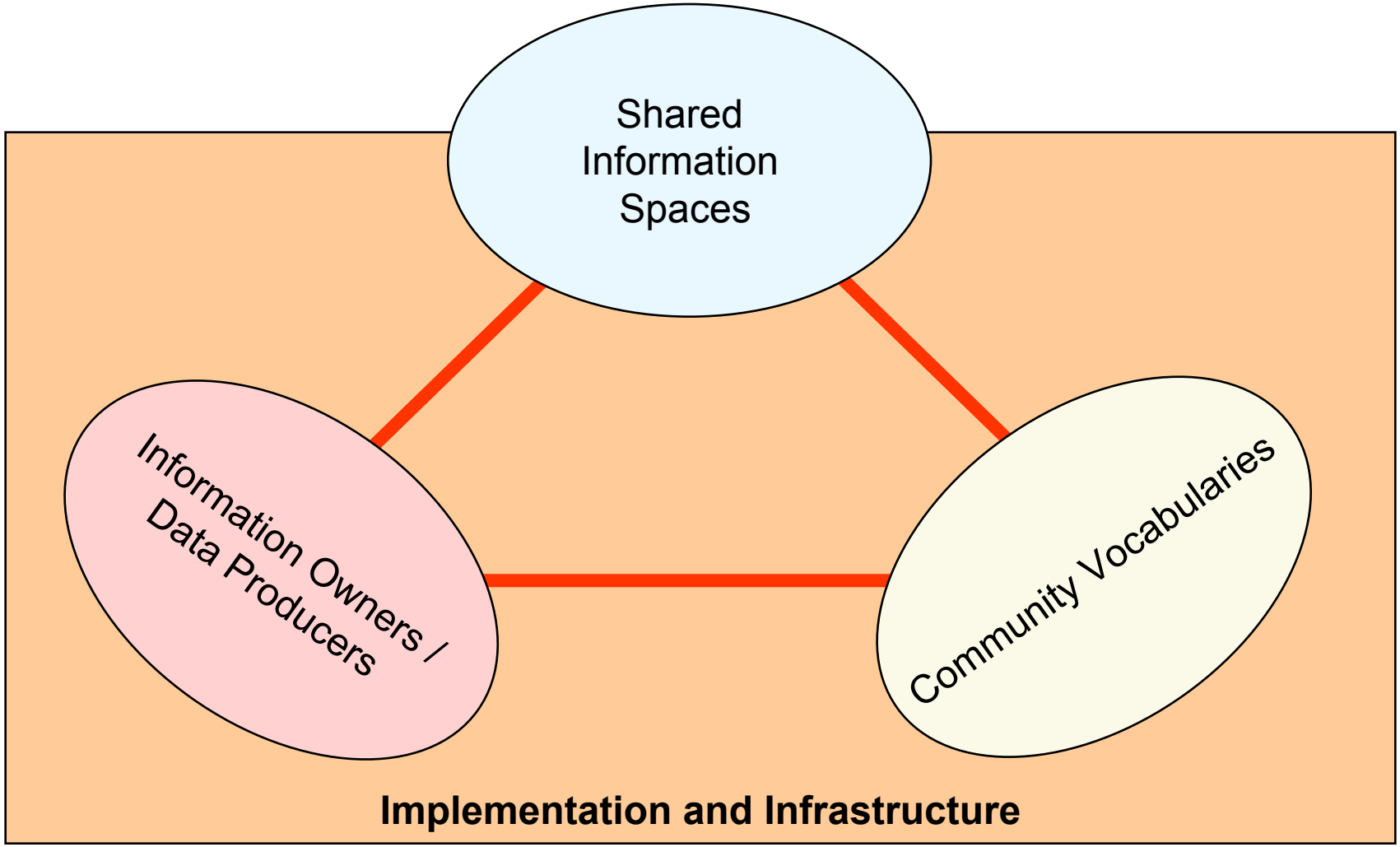
The Air Force Chief Information Officer (AF CIO) is responsible for establishing AF information and data management policies. Establishment of Air Force COIs is delegated to the Enterprise Architecture Integration Council (EAIC). The implementing roles and responsibilities of information owners and COIs are included in the attachment to this memorandum, together with an initial list of AF COIs. Information owners will be designated separately. Air Force MAJCOMS and Functional Community leads will execute their data producer responsibilities, coordinating through COIs as described in that attachment.

John P. Jumper
General, USAF
Chief of Staff

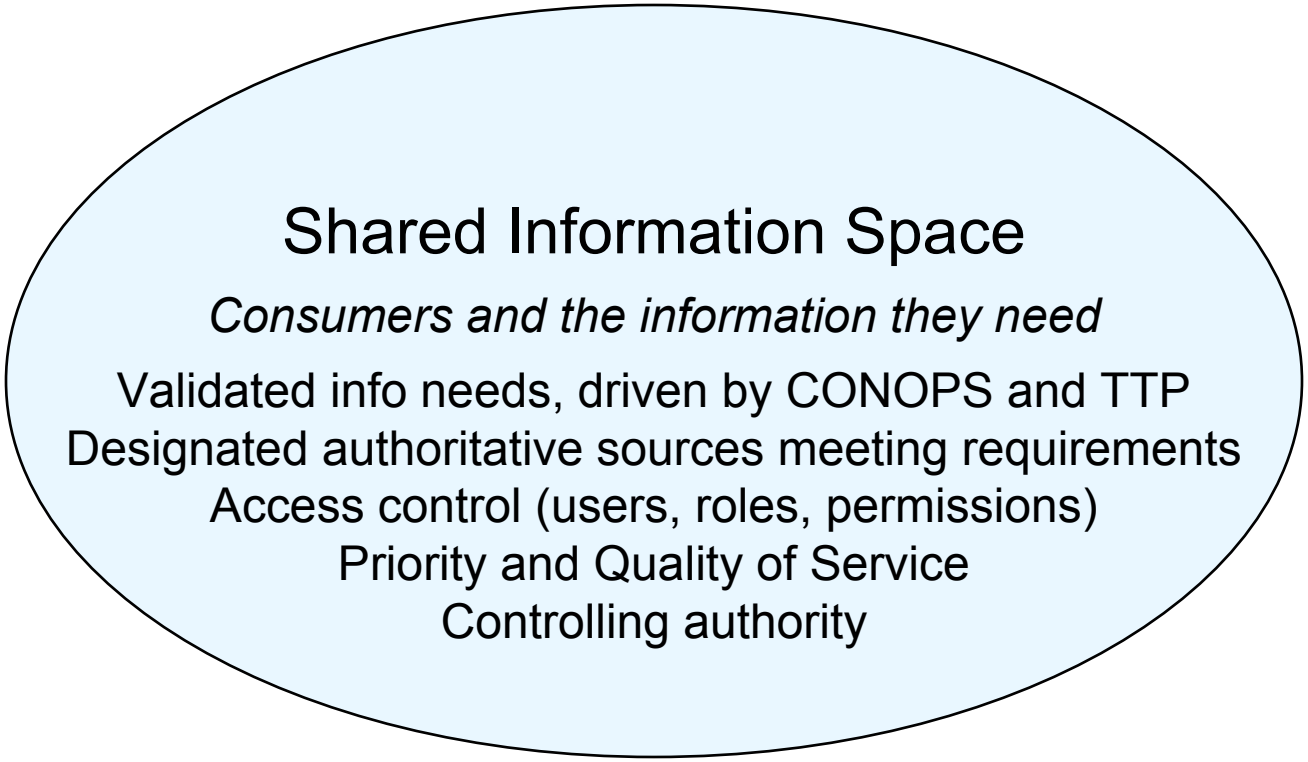
James G. Roche
Secretary of the Air Force

Attachment:
Roles and Responsibilities

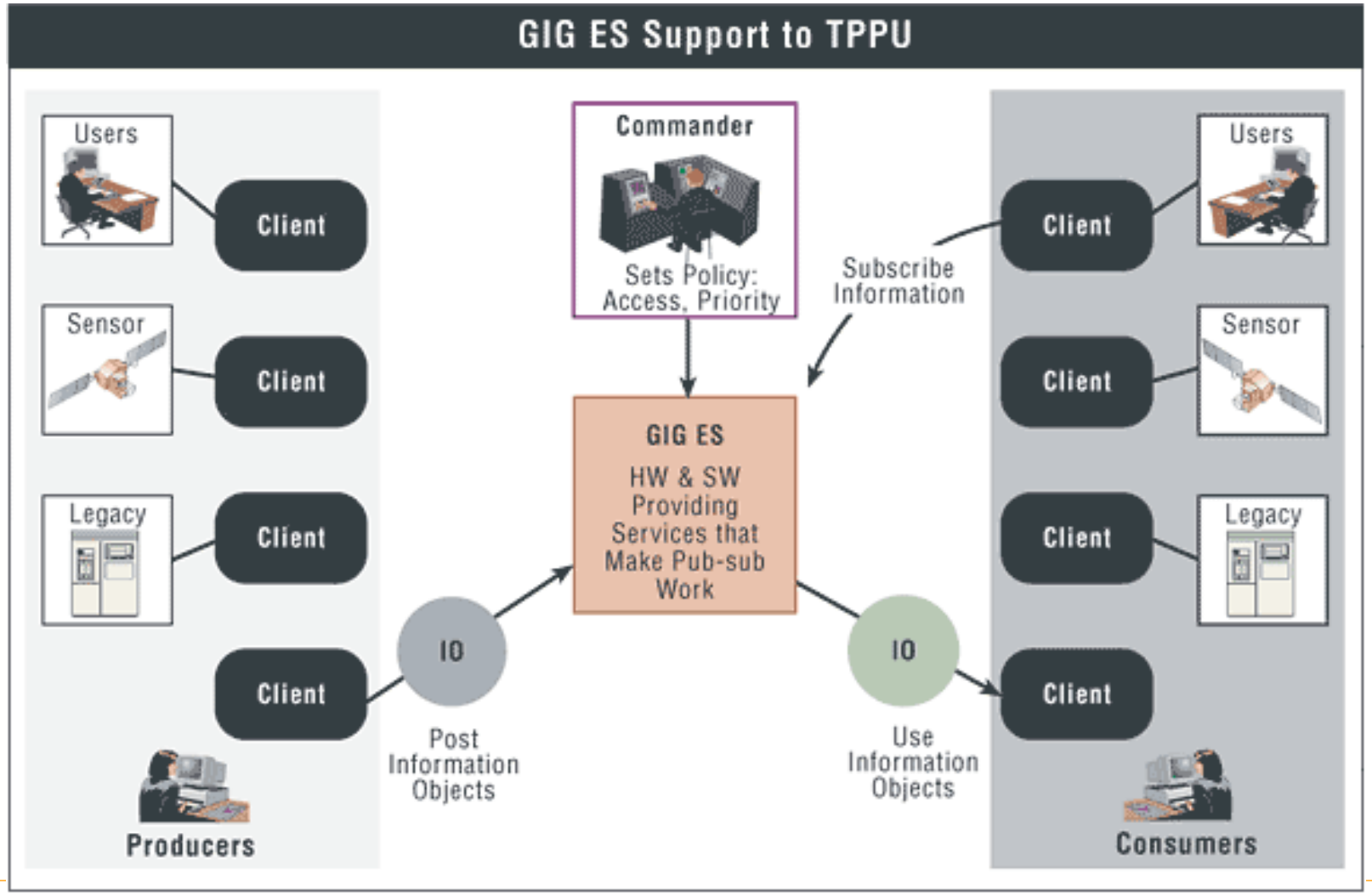
AF Information & Data Mgt. Strategy: The "Triangle Foundation"



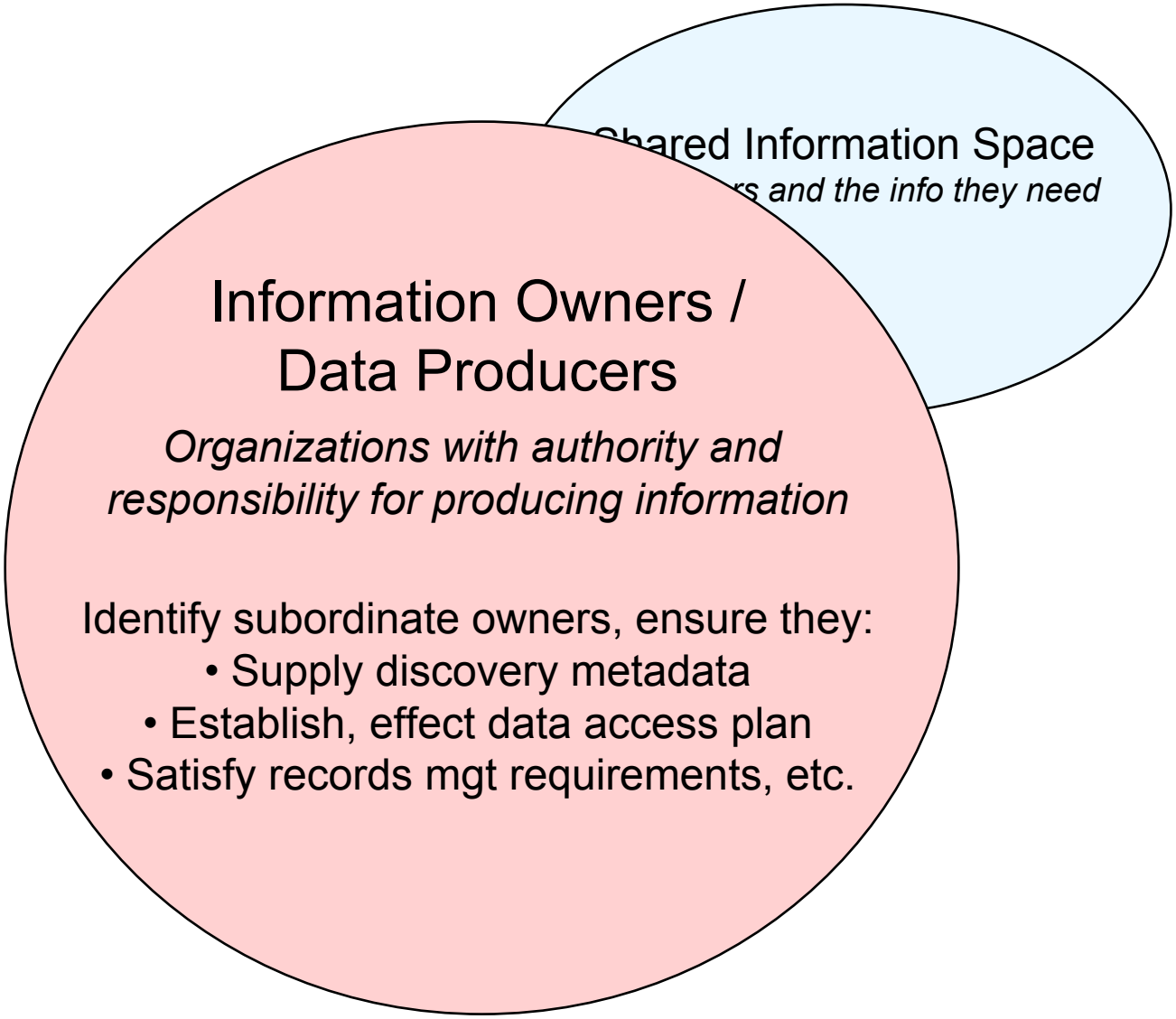
Shared Information Spaces



Task, Post, Process, Use (TPPU): An Infospace and Controlling Authority

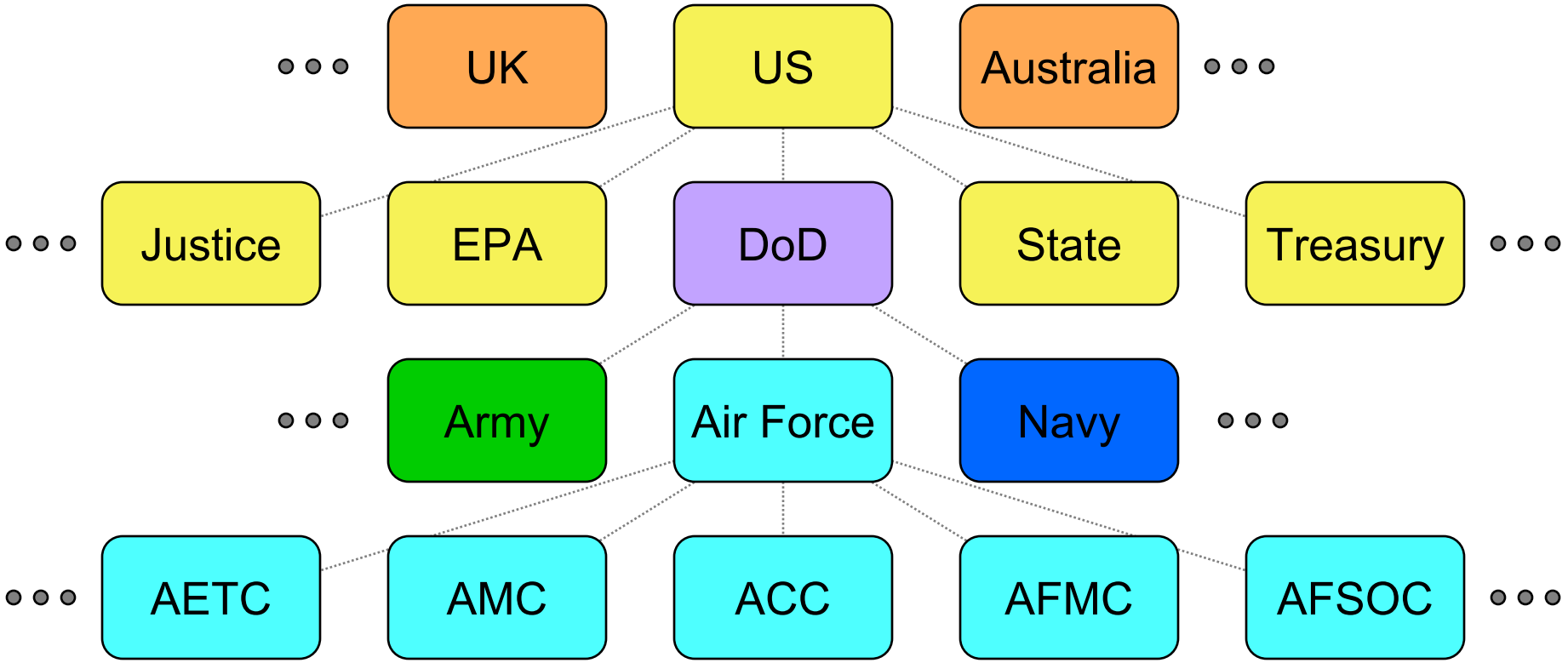


Information Owners

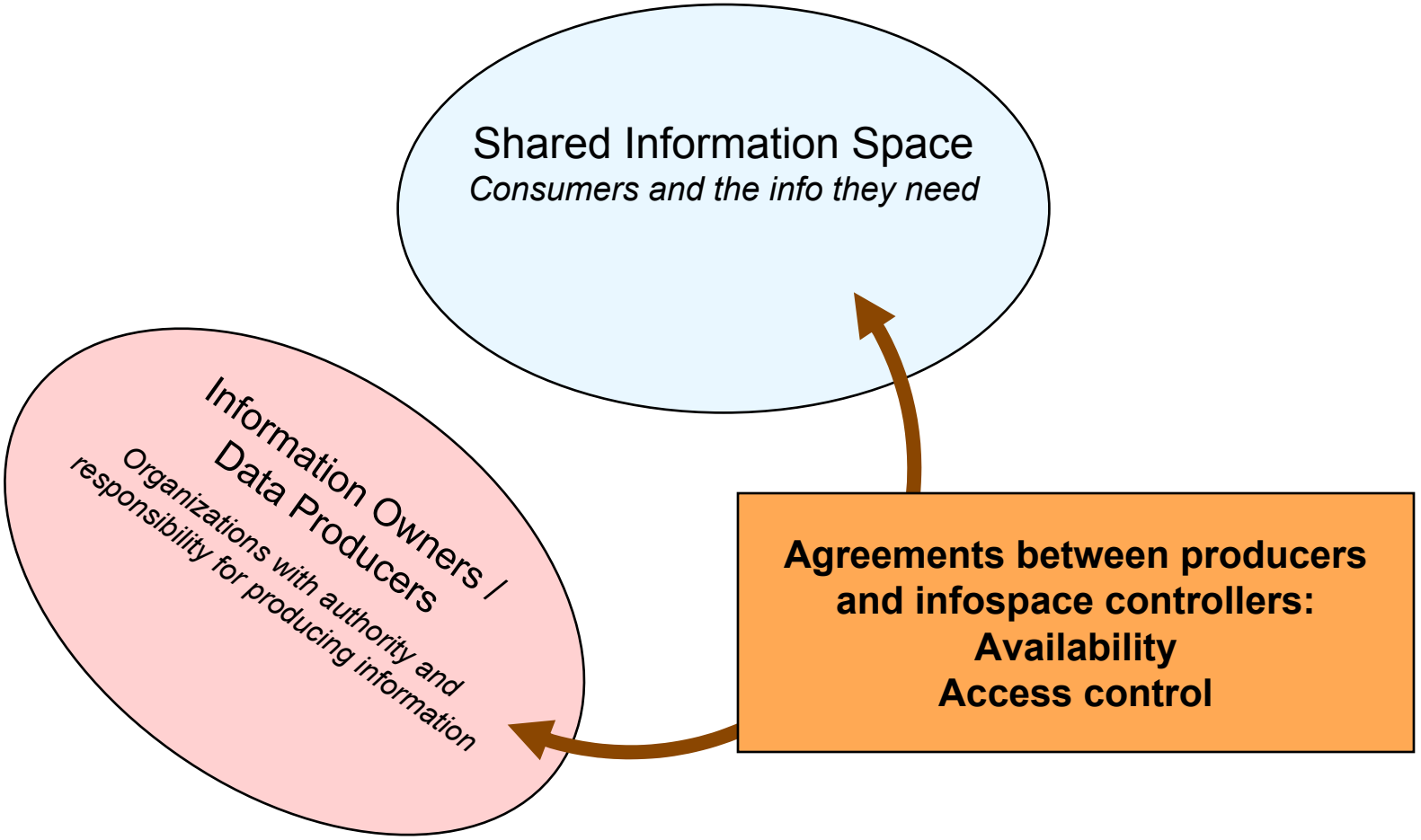


The Enterprise: Subordinates and Partners

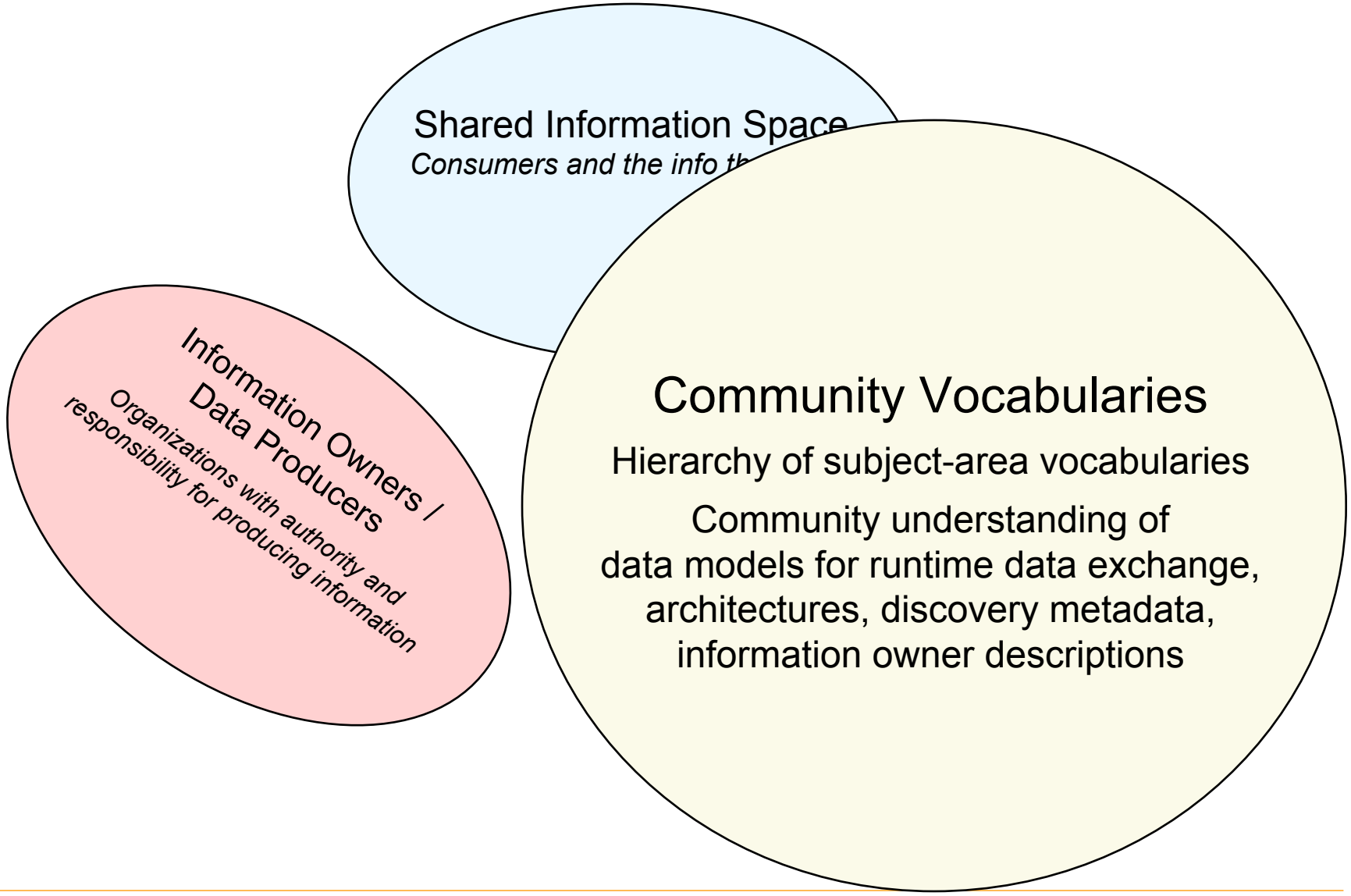
You can't just say "all information is owned by the enterprise"



Data Service Agreements Are Required



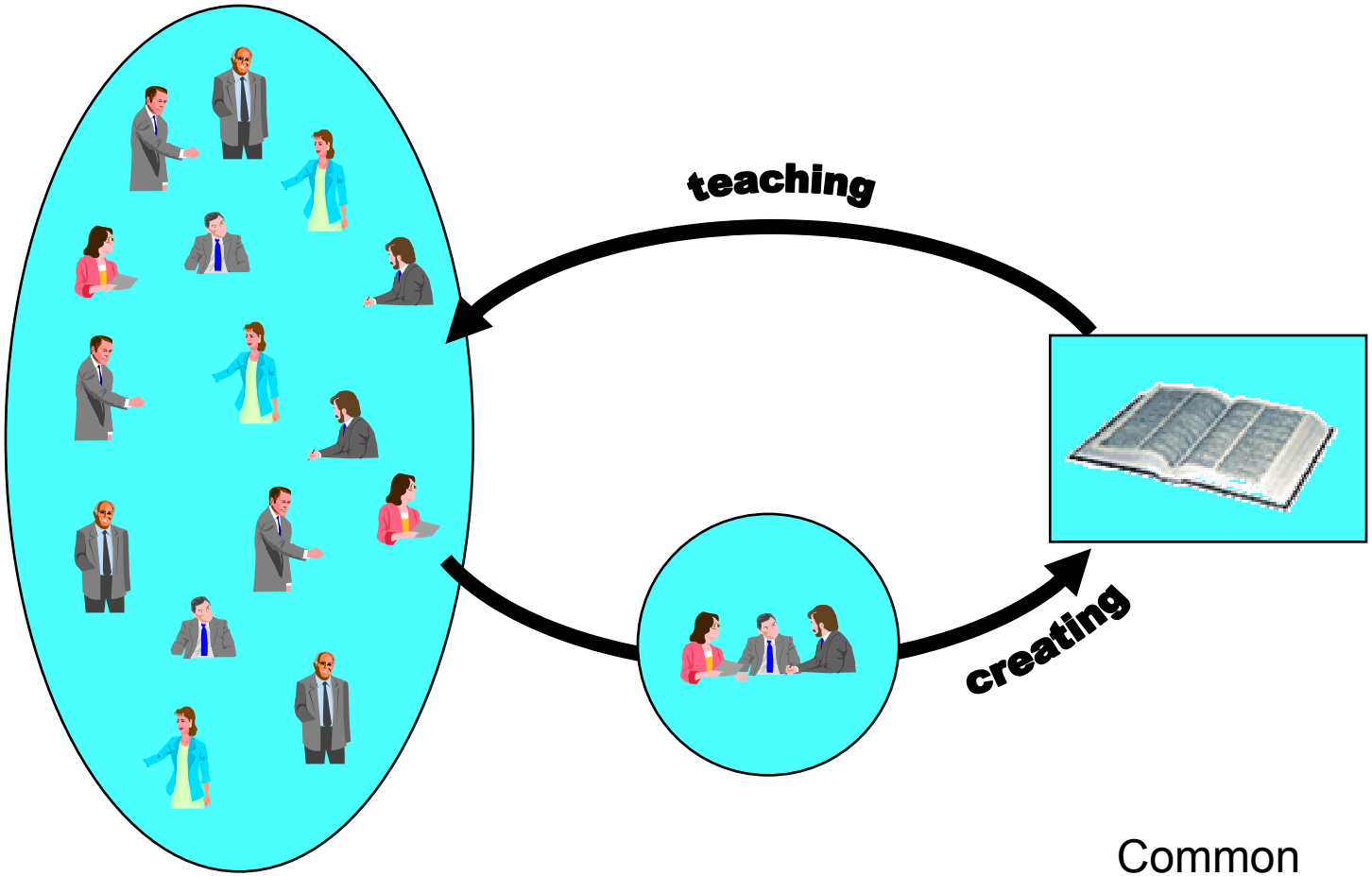
Community Vocabularies



Community Vocabularies

- Why “community vocabularies”? Why not talk about
  - Data models
  - Data elements
  - Schemas
  - Taxonomies
  - Ontologies
  - Other formal methods to capture / share semantics
- We need shared semantics for several purposes
  - Data definitions for programmers and users
  - Metadata “tags” for discovery
  - Information description in architecture products
- We use “vocabulary” to subsume all of those formalisms for all of those purposes

Vocabulary Is Knowledge



Community of interest

Data panel

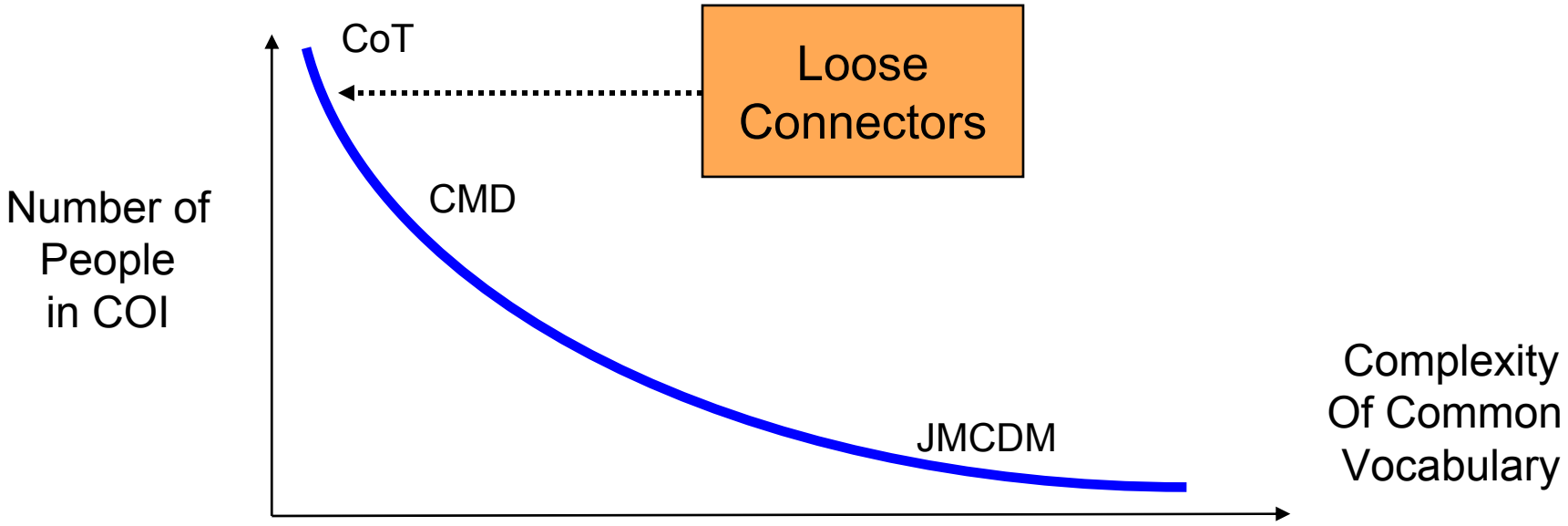
Common vocabulary (ontology, etc.)

The Power Law of Common Vocabularies

- **We should expect a range between:**
  - **A small number of broad, shallow vocabularies: few definitions, understood by many people**
  - **A larger number of narrow, deep vocabularies: many definitions, understood by few people**

CoT

JMCDM
Joint METOC Conceptual Data Model

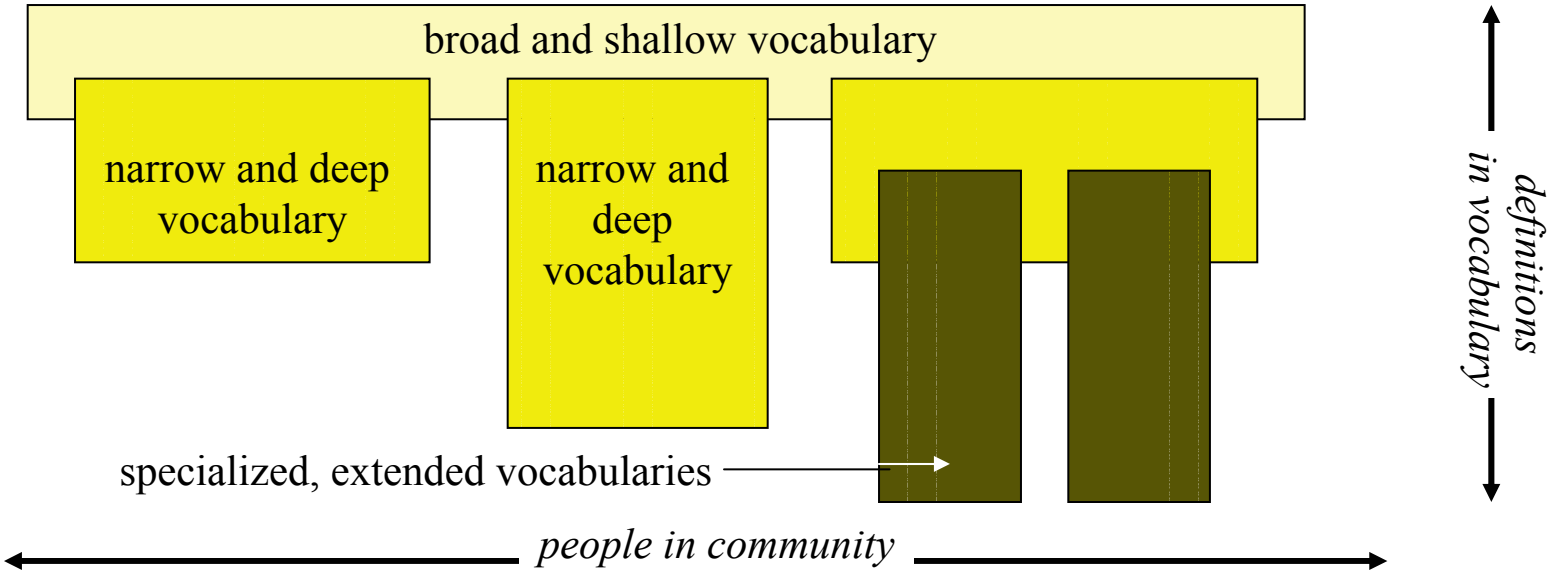


The Power Law of Common Vocabularies

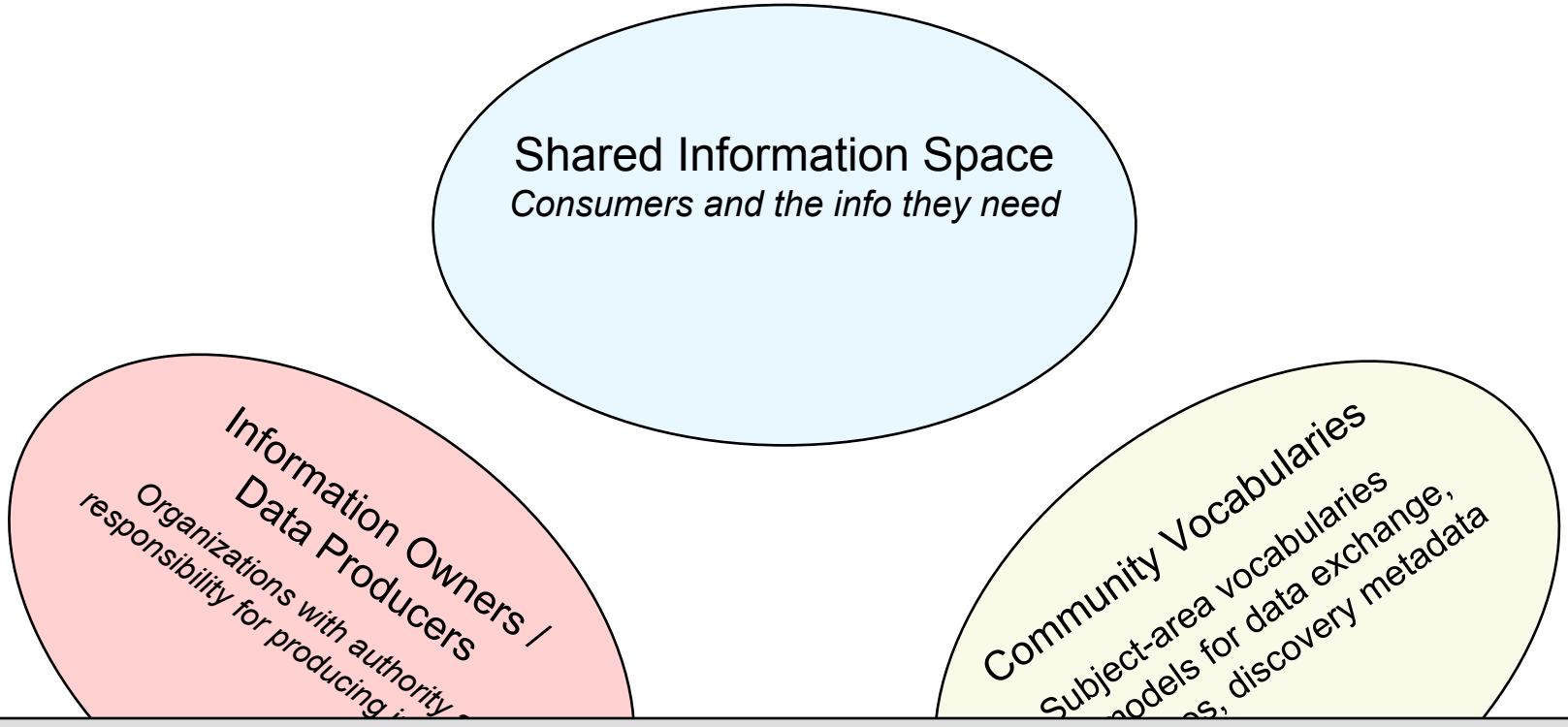
- **We should expect a range between:**
  - **A small number of broad, shallow vocabularies: few definitions, understood by many people**
  - **A larger number of narrow, deep vocabularies: many definitions, understood by few people**
  - **Eventually, deep vocabularies will extend and specialize the shallow vocabularies**

CoT

JMCDM
Joint METOC Conceptual Data Model



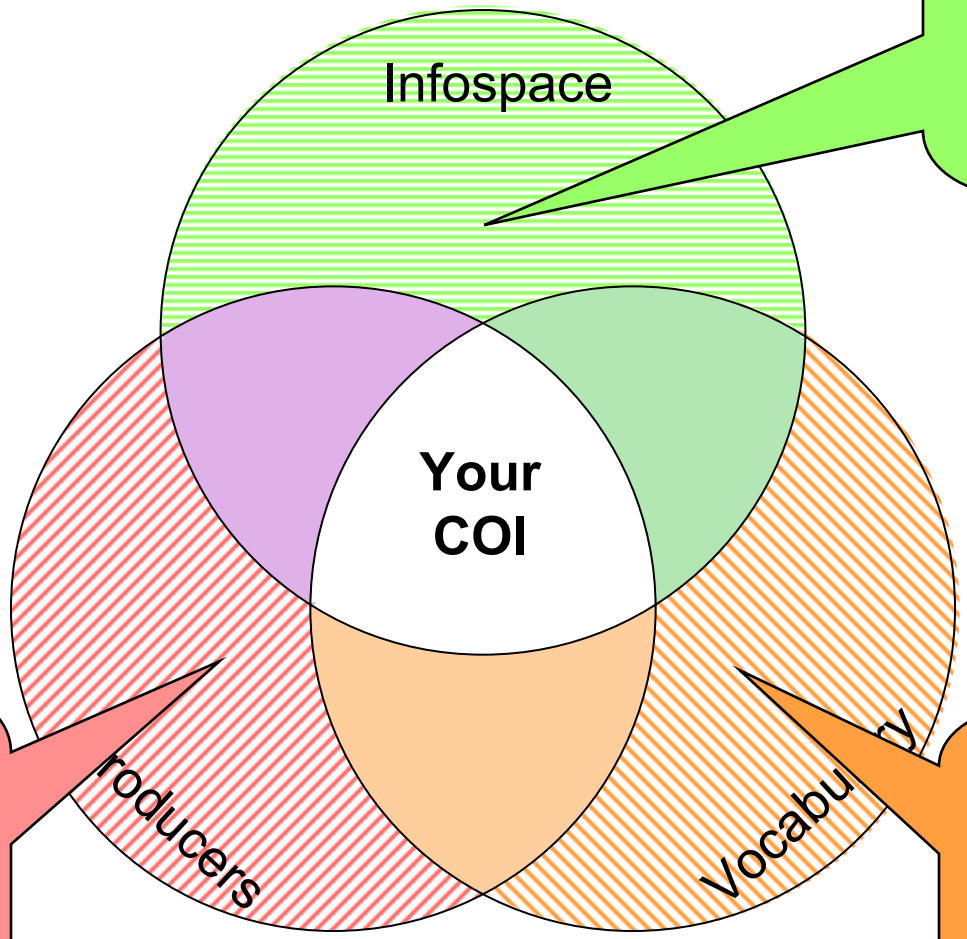
Implementation Infrastructure



Implementation and Infrastructure

Data access implementation (warehouse, pub/sub, web service, portal, etc.)
Physically located at producers, consumers, or someplace in between.
Core services for identity, authentication, etc.

Expanding To The Enterprise

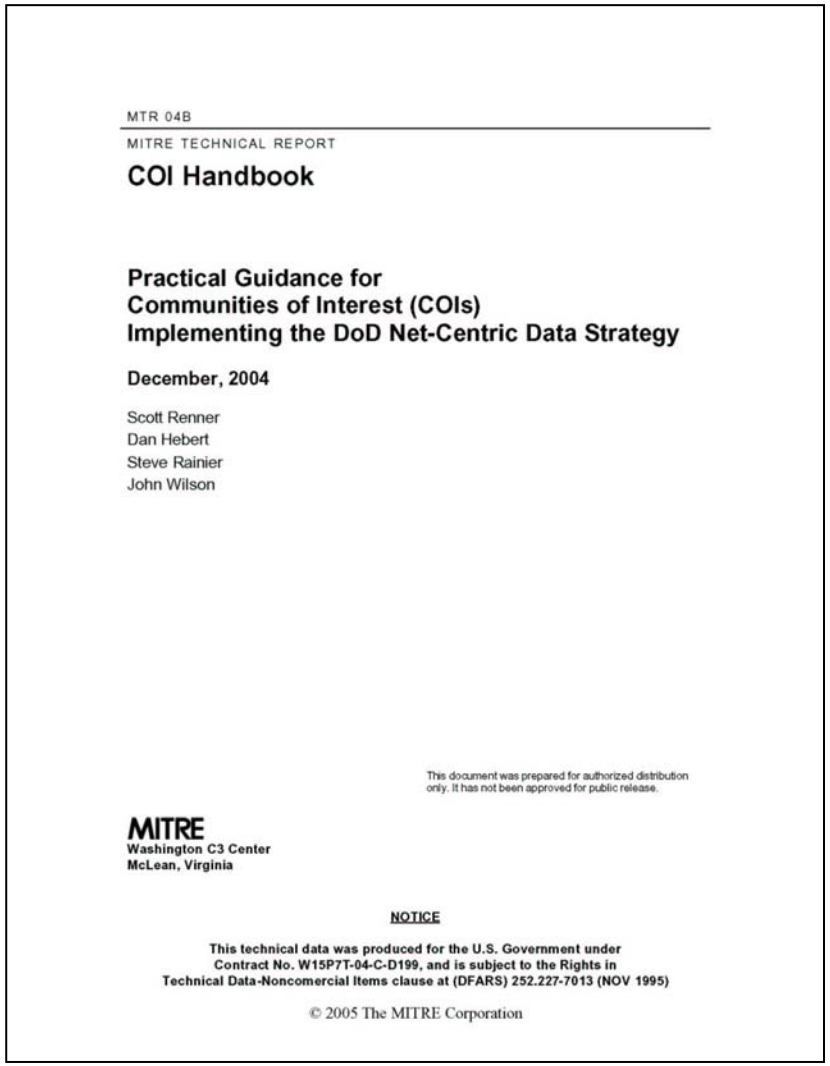


The infospace will include information from other producers, following other vocabularies

Producers will post to other infospaces, using other community vocabularies

Other producers and other infospaces will use the community vocabulary

The COI Handbook



MTR 04B

MITRE TECHNICAL REPORT

COI Handbook

Practical Guidance for
Communities of Interest (COIs)
Implementing the DoD Net-Centric Data Strategy

December, 2004

Scott Renner
Dan Hebert
Steve Rainier
John Wilson

This document was prepared for authorized distribution only. It has not been approved for public release.

MITRE
Washington C3 Center
McLean, Virginia

NOTICE

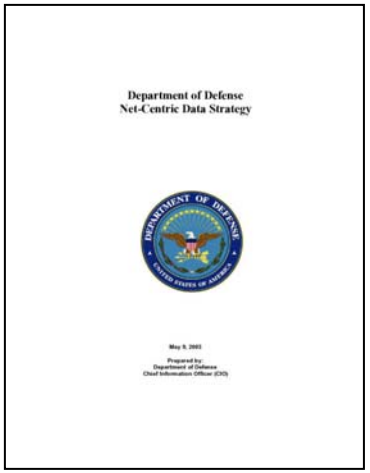
This technical data was produced for the U.S. Government under Contract No. W15P7T-04-C-D199, and is subject to the Rights in Technical Data-Noncomercial Items clause at (DFARS) 252.227-7013 (NOV 1995)

© 2005 The MITRE Corporation

- Tasks and Responsibilities
  - Shared Vocabulary
  - Shared Information Space
  - Information Owners and Data Producers
- Implementation Guidance
  - Exploration Spiral
  - Implementation Spiral
  - Operations Spiral
- Case Study

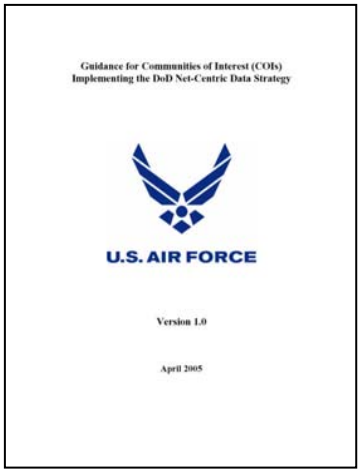
COI Handbook In Context

COI Handbook



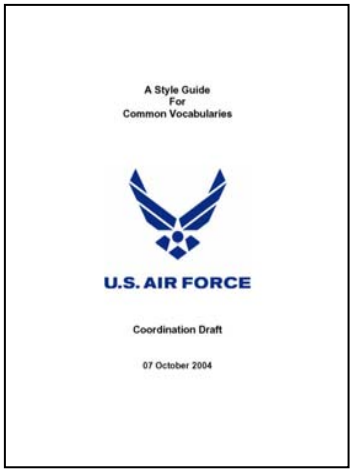
Net-Centric Data Strategy

AF I&DMS

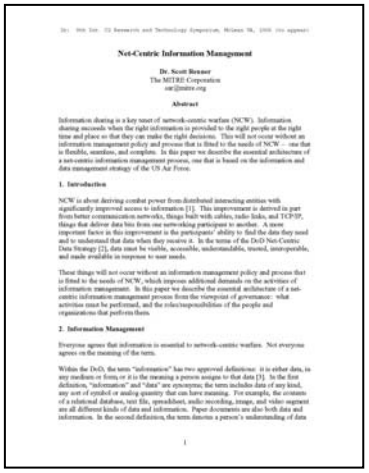


AF COI Guidance

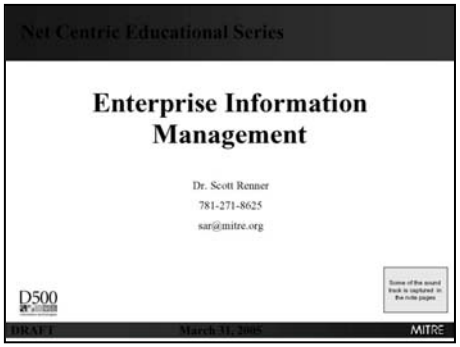
COI Glossary



Common Vocab. Style Guide



Net-Centric IM



Enterprise IM

Summary

- Purpose of information management
  - Right information to right person at right time and place
  - Ensure that information exists, is discoverable and understandable
- Elements of net-centric IM
  - Shared information spaces
  - Info owners / data producers
  - Community vocabularies
  - Implementation and infrastructure